ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation

Kasper Jermiin Knudsen[1,2,3§], Matilda Rehn[1,2,3§], Marie Sigurd Hasemann[1,2,3], Nicolas Rapin[1,2,3,4], Frederik Otzen Bagger[1,2,3,4], Ewa Ohlsson[1,2,3], Anton Willer[1,2,3], Anne-Katrine Frank[1,2,3], Elisabeth Søndergaard[1,2,3], Johan Jendholm[1,2,3], Lina Thorén[1,2,3], Julie Lee[1,2,3], Justyna Rak[5], Kim Theilgaard-Mönch[1,2,5,6*], Bo Torben Porse[1,2,3,*]

Supplemental Material contains:

- Nine supplemental tables, Table S1–S9 containing data related to gene expression profiling and bioinformatics analyses, mainly discussed in connection to Figure 5 (provided as separate files)
- Four Supplemental Figures: Figure S1 related to Figure 1
  Figure S2 related to Figure 2
  Figure S3 related to Figure 2 and 3
  Figure S4 related to Figure 4
- Legends to Supplemental Tables S1–S9
- Supplemental Material and Methods
- Supplemental References

Supplemental Figures

Knudsen_Figure S1

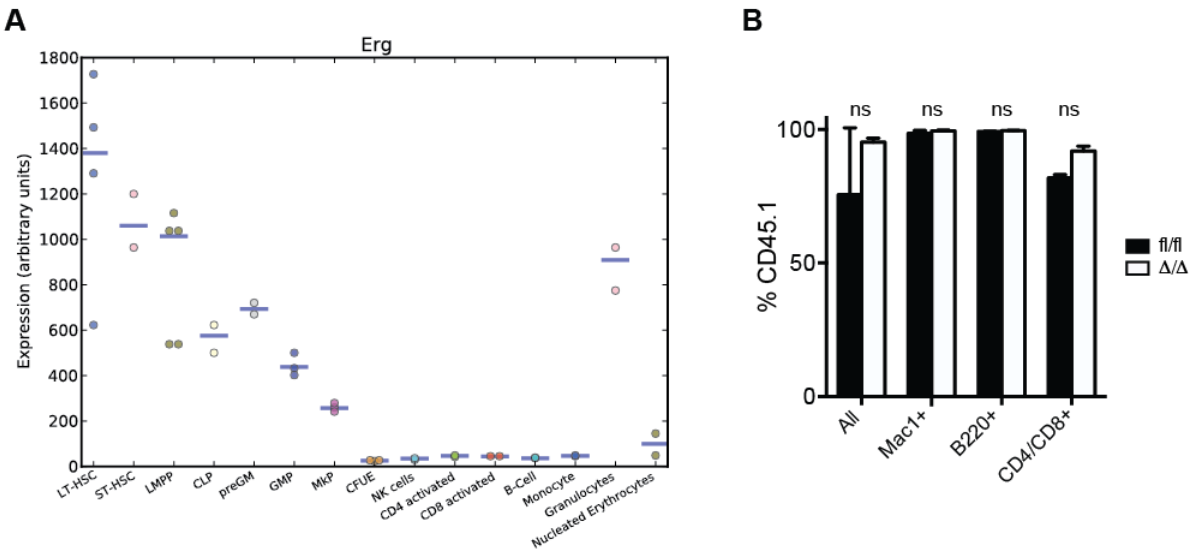

Figure S1. *Erg* is highly expressed in HSC/HPCs and the HSC defect upon *Erg* loss is cell intrinsic. (**A**) *Erg* expression in different hematopoietic cell populations derived from the HemaExplorer website (http://servers.binf.ku.dk/hemaexplorer/). (**B**) Analysis of the PB in irradiated CD45.2 $Erg^{fl/fl}$ (n=5) and $Erg^{\Delta/\Delta}$ (n=7) mice, 16 weeks after transplantation of WT BM (CD45.1). These data illustrate the lack of hematopoietic defects in pIpC injected $Erg^{fl/fl}$; *Mx1Cre* if rescued with WT BM cells. Data are represented as mean + SD. ns; not significant.

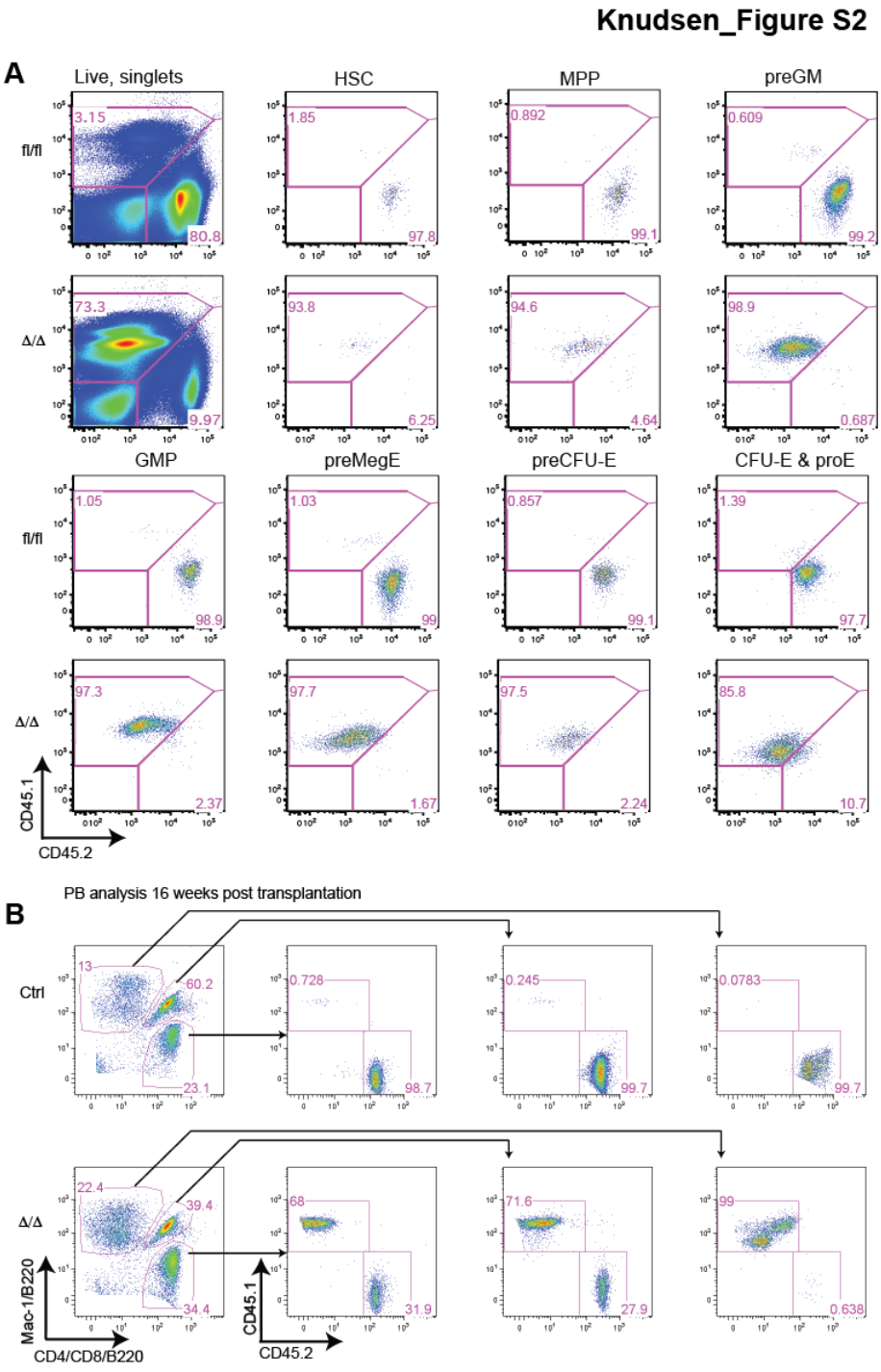

Figure S2. *Erg* deficient HSCs are efficiently outcompeted by WT HSCs. (**A**) Representative FACS data of the results presented in Figure 2D. FACS analysis of BM from CD45.1 recipients, 20 weeks after transplantation of $Erg^{fl/fl}$ or $Erg^{fl/fl}$*; Mx1Cre* (CD45.2) BM in a 10:1 ratio with competitor BM. (**B**) Representative FACS data of the results presented in Figure 2I. FACS analysis of PB from non-conditioned $Erg^{fl/fl}$ and $Erg^{fl/fl}$*; Mx1Cre* (CD45.2) recipients of WT (CD45.1) BM. Mice were injected with pIpC prior to transplantation and analyzed 16 weeks post-transplant.

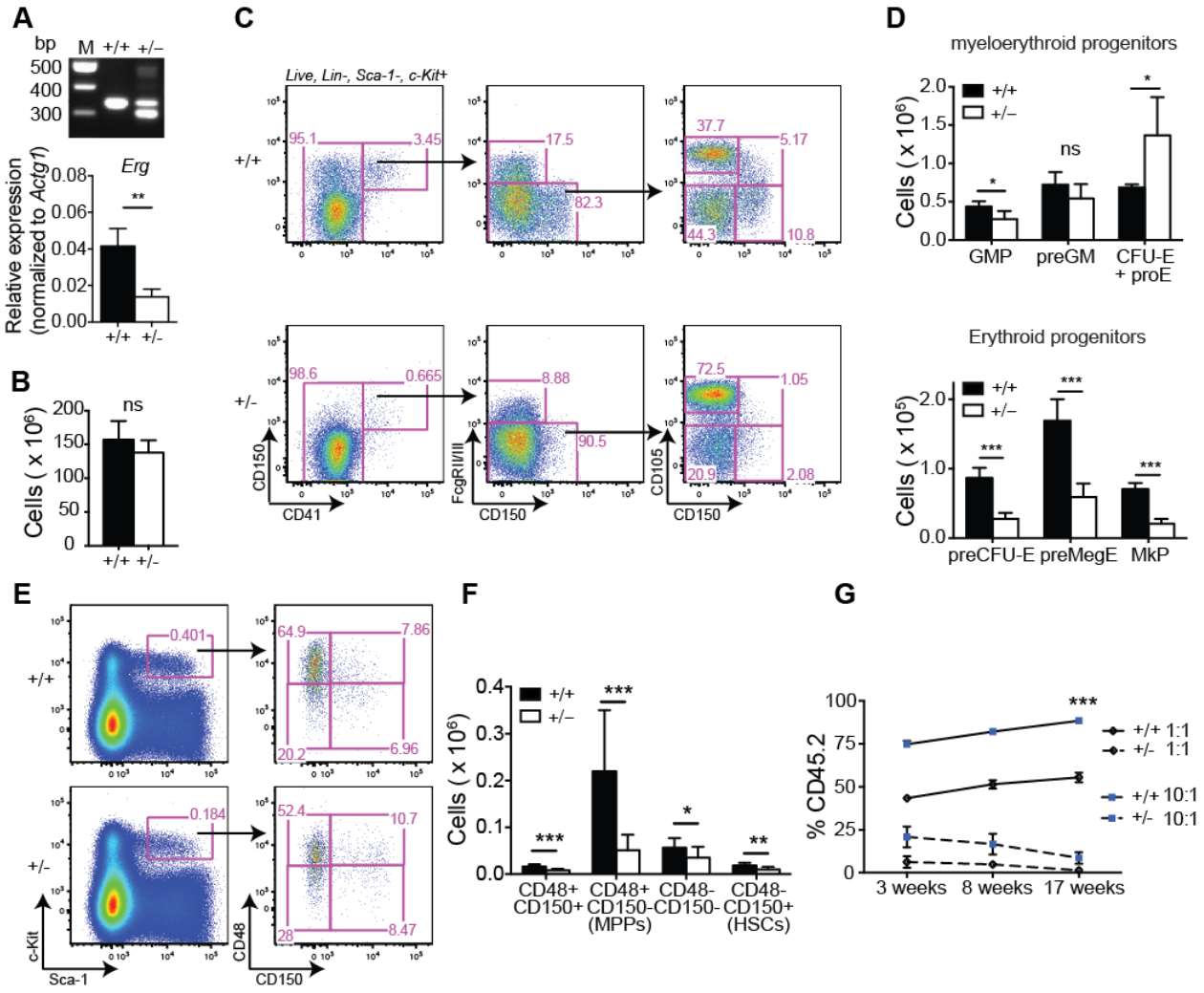

Figure S3. **The effect of ERG on HSC maintenance is dose-dependent.** (**A**) Genotyping of *Erg*$^{+/+}$ and *Erg*$^{+/-}$ BM cells. (**B**) Cell numbers in BM (2x *femur, tibia* and *ilium*) of *Erg*$^{+/+}$ (n=8) and *Erg*$^{+/-}$ (n=8) mice. (**C**) Distribution of myeloid progenitors in *Erg*$^{+/+}$and *Erg*$^{+/-}$ BMs (**D**) Quantification of the data in (**C**); *Erg*$^{+/+}$ (n=4) and *Erg*$^{+/-}$ (n=4). (**E**) Distribution of stem and progenitor cells in *Erg*$^{+/+}$ and *Erg*$^{+/-}$ BMs. (**F**) Quantification of the data in (**E**); *Erg*$^{+/+}$ (n=12) and *Erg*$^{+/-}$ (n=13). (**G**) Analysis of PB from mice transplanted with *Erg*$^{+/+}$ (n=5) and *Erg*$^{+/-}$ (n=5) BM in a 1:1 or 10:1 ratio with competitor BM. Data are represented as mean +SD. *, P < 0.05; **, P < 0.01; ***, P < 0.001. ns; not significant; M;marker, GMP; granulocyte-macrophage progenitor, preGM; pre-granulocyte-macrophage, CFU-E; erythroid colony-forming-unit, pro-E; pro-erythroblast, preCFU-E; pre erythroid colony-forming unit, preMegE; pre megakaryoblast-erythroid progenitor, MkP; megakaryocyte progenitor.

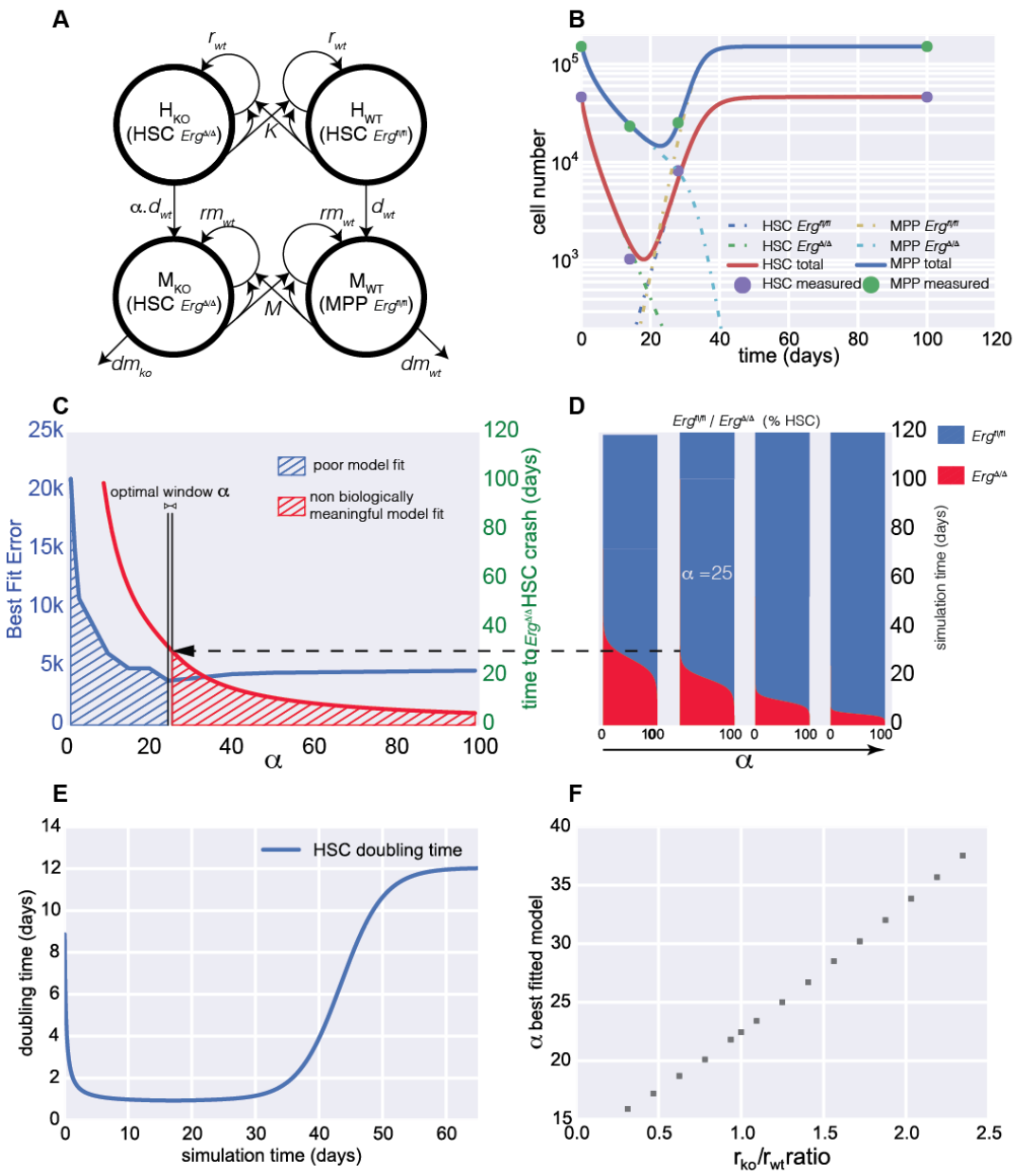

Figure S4. **Modelling HSC loss in pIpC-injected *Erg*$^{fl/fl}$; *Mx1Cre* mice.** (**A**) Flowchart of the ordinary differential equation model presented in equation 1.10. (**B**) Best model fit for cell populations in *Erg*$^{\Delta/\Delta}$ BM, obtained with parameters from *Modelling 3*. Experimental data points are represented by dots; total HSC (HSC *Erg*$^{\Delta/\Delta}$ and HSC *Erg*$^{fl/fl}$) and MPP (MPP *Erg*$^{\Delta/\Delta}$ and MPP *Erg*$^{fl/fl}$) populations are represented by thick lines. (**C**) Determination of a best parameter value window for α. As the parameter value is increased, the model is fitted for all other parameters and the squared error is reported. A minimal error value is obtained for α = 24.31 (*Modelling 3*). At the same time we report the time point where *Erg*$^{\Delta/\Delta}$ HSCs disappear from the system and match the results with experimental data. A small window is found where the model fit error is minimal and where the model agrees with experimental data. (**D**) Proportion of HSC *Erg*$^{\Delta/\Delta}$ and HSC *Erg*$^{fl/fl}$ versus simulation time, as α is increased. (**E**) Doubling time of HSCs as a function of simulation time; for parameter values used, see *Modelling 3*. (**F**) Variation in differentiation (α) as a function of proliferation rate ratio between *Erg*$^{\Delta/\Delta}$ HSCs and *Erg*$^{fl/fl}$ HSCs; other parameter values taken from *Modelling 3*.

Legends to Supplemental Tables

Table S1. Gene expression profiles of $Erg^{+/-}$ (n=3) *vs*. $Erg^{+/+}$ (n=3) HSCs (LSK, CD150+, CD48-).

Table S2. Gene expression profiles of $Erg^{\Delta/\Delta}$ (n=3) *vs*. $Erg^{fl/fl}$ (n=3) MPPs (LSK, CD150-).

Table S3. Genes downregulated by ≥ 1.5 fold change in $Erg^{+/-}$ *vs*. $Erg^{+/+}$ which have an ERG-bound region within +/- 800 bp from TSS.

Table S4. Genes upregulated by ≥ 1.5 fold change in $Erg^{+/-}$ *vs*. $Erg^{+/+}$ which have an ERG-bound region within +/- 800 bp from TSS.

Table S5. Gene ontology analysis on ERG-bound downregulated genes (from Table S3).

Table S6. Gene ontology analysis on ERG-bound upregulated genes (from Table S4).

Table S7. DREME motif analysis of all ERG-bound region in HPC7 cells.

Table S8. Occurrence of selected motifs from Table S7 in ERG-bound deregulated genes in HSCs.

Table S9. Gene ontology analysis on ERG-bound deregulated genes (from Table S8).

Supplemental Material and Methods

Mouse work

A lambda phage clone containing exon 11 (encoding the DNA binding domain) of *Erg* was retrieved using the REC screening system (Zhang et al. 2002). LoxP sites were subsequently introduced on each side of exon 11 using homologous recombination in *E.coli* and the resulting targeting construct was introduced into E14.1 ES cells by electroporation. Blastocyst injections and breeding for germ line transmission was done as described previously (Weischenfeldt et al. 2008). The Frt-flanked Neostop cassette was subsequently removed by crossing with a Flp deleter line (Dymecki 1996) resulting in mice heterozygous for the Erg^{flox} (Erg^{fl}) allele. In order to generate mice heterozygous for the *Erg* allele, we crossed $Erg^{fl/+}$ with a general Cre deleter mouse (Schwenk et al. 1995). Mice with the Erg^{fl} allele were bred with mice harbouring the *Mx1-Cre* allele (Kuhn et al. 1995), the *R26-CreER* (Ventura et al. 2007) and the *CD2iCre* allele (de Boer et al. 2003), and subsequently backcrossed to a C57BL/6 background. All mice have been backcrossed for at least six generations. Animals were housed according to institutional guidelines at the University of Copenhagen and experiments were performed under permission from the Danish Animal Research Ethical Committee.

Excision of the Erg^{fl} allele was achieved by subjecting 10–12 weeks old $Erg^{fl/fl}$*; Mx1Cre* and $Erg^{fl/fl}$ mice to 3–5 injections (two days interval) with 200 µl polyinosinic-polycytidylic acid (pIpC) (1.5 mg/ml, GE-Healthcare).

Mice were genotyped using a three-primer PCR (ATCATGACAATAAGCCGGGT; AACCAGAACGGTGGTAGTCT; GTGGTGGCTCCTTAAGGGTC) resulting in products of the following sizes: 340 bp (Erg^{+}) 397 bp (Erg^{fl}) and 302 bp (Erg^{Δ}).

Ki67 and cell death analyses

BM cells from *Erg*$^{+/-}$ and *Erg*$^{+/+}$ mice were stained for HSCs (LSK, CD150+, CD48-) and fixed in 4% paraformaldehyde (Sigma-Aldrich) for 10 min at room temperature (RT) followed by permeabilisation with 0.1% saponin (Sigma-Aldrich) in PBS+3% FCS for 45 min at RT. Next, cells were stained with anti-Ki67 (BD) for 30 min at RT and with 0.5 μg/ml DAPI (Molecular Probes) and analyzed on a LSRII (BD).

ROS staining

BM cells from *Erg*$^{+/-}$ and *Erg*$^{+/+}$ mice were enriched for c-Kit and stained for HSCs (LSK, CD150+, CD48-) prior to staining with CM-H2DCFDA (Molecular Probes) according to manufacturers protocol.

Homing assays

Homing assays were performed with either c-Kit-enriched BM cells from *Erg*$^{fl/fl}$ or *Erg*$^{\Delta/\Delta}$ mice or HSCs (LSK, CD150+, CD48-) derived from *Erg*$^{+/-}$ or *Erg*$^{+/+}$ mice. The cells were stained with carboxyfluorescein-diacetate-succinimidyl-ester (CFSE) (CellTrace^{TM}, Invitrogen) according to manufacturers protocol. Briefly, c-Kit-enriched BM or 10 000 FACS sorted HSCs were incubated at 37° C for 10 min with 10 μM CFSE in 200 μl PBS+3% FCS. Next, cells were washed three times as follows: one ml ice cold PBS+3% FCS was added followed by five min incubation on ice and centrifugation. Finally, HSCs (3000 cells/mouse) or c-Kit-enriched BM ($7x10^5$cells/mouse) were resuspended in PBS+3% FCS and transplanted by tail vein injection into irradiated (900 cGy) Ly-5.1 (CD45.1) mice. Three (c-Kit-enriched BM) or twelve (HSCs) hours later, BM and spleen from recipients were isolated, enriched for c-Kit+ cells as described above, and analyzed on a LSRII for CFSE+ cells. Homing efficiency was defined as recovered cells in the BM divided by the amount of

transplanted cells.

Colony assays

Splenocytes or FACS-sorted hematopoietic populations (HSC: LSK CD150+; MPP: LSK CD150-; preGM: Lin- c-Kit+ Sca-1- CD41- FcgRII/III- CD150-; GMP Lin- c-Kit+ Sca1- CD41- CD150- FcgRII/III+) were plated in M3434 semisolid methylcellulose media (Stemcell Technologies) supplemented with penicillin/streptomycin (100 U/100 μg/ml, Invitrogen) according to manufacturers protocol. On day 10, the colonies were scored by inverted microscopy and individual colonies were picked and assessed for recombination.

Single cell HSC cultures

Single HSCs (LSK, CD150+, CD48-) from *Erg*$^{fl/fl}$ or *Erg*$^{fl/fl}$; *R26-CreER* were sorted directly to round-bottom 96-well plates and cultured in serum-free medium (StemSpan®SFEM, Stemcell Technologies) supplemented with 100 ng/ml murine SCF (mSCF, R&D Systems), 100 ng/ml human Flt3L (hFlt3L, R&D Systems), 100 ng/ml human TPO (hTPO, Peprotech) and 1 μM (Z)-4-hydroxytamoxifen (4-OHT, Sigma-Aldrich). Fresh media was added on day two, four and six. Each well was examined by microscopy and cell numbers were recorded on day eight.

LTC-IC assays

FBMD-1 stromal cells (Ploemacher et al. 1989; Breems et al. 1994) were cultured in FBMD-1 medium consisting of IMDM (Invitrogen), 5% horse serum (Invitrogen), 10% FCS, 0.1 mM beta-mercaptoethanol (Sigma-Aldrich) 100 U/100 μg/ml penicillin/streptomycin and 10 μM hydrocortisone (Sigma-Aldrich).

Individual $Erg^{fl/fl}$ or $Erg^{\Delta/\Delta}$ HSCs were sorted one week after *Erg* excision into flat-bottomed surface-treated 96-well plates containing a confluent layer of FBMD-1 cells and cultured for five weeks in FBMD-1 co-culture media (IMDM, 20% horse serum, 0.1 mM beta-mercaptoethanol, 100 U/100 μg/ml penicillin/streptomycin, 10 μM hydrocortisone) with weekly medium changes. After five weeks, media was replaced with 150 μl M3434 semi-solid medium (Stemcell Technologies) and incubated for 10 days before positive wells were scored by microscopy. LTC-IC frequencies were defined as number of positive wells divided by number of seeded wells multiplied by 100.

In vitro HSC adhesion assays

FBMD-1 stromal cells were seeded in 96-well plates and cultured for a week. To obtain GFP-labeled HSCs, $Erg^{+/-}$ mice were intercrossed with *B10q-GFP* mice, the latter of which expresses GFP ubiquitously (Okabe et al. 1997). GFP+ $Erg^{+/-}$ and $Erg^{+/+}$ HSCs (LSK CD150+ CD48-) were sorted directly into 96-well plates (100 HSCs/well) in IMDM, 10% FCS. HSCs were allowed to adhere for one hour at 37° C after which the plate was slowly immersed in PBS pre-heated to 37° C. A PBS-filled 96-well plate was placed on top and fixed with a steel frame so that individual wells from the two plates were in contact. Next, the two plates were inverted 180°, still in the PBS container. The inverted plates were incubated for one h at 37° C to allow for non-adhering HSCs to fall down by gravitational forces. The upper 96-well plate, containing the FBMD-1 stroma, was then removed from the PBS container and GFP+ HSCs still attaching to the stromal cells were quantified by fluorescence microscopy.

Primers used for qRT-PCR analysis

Gene	Forward primer	Reverse primer
Actg1	CTCTTCCAGCCTTCCTTCCT	TGCTAGGGCTGTGATCTCCT
Ccnb1	TCTTCTCGAATCGGGGAAC	GACCTTGGCCTTATTTTCTGC
Cks2	TCGATGAGCACTACGAGTACC	CCATCCTAGACTCTGTTGGACAC
Dkc1	AAAGACCGGAAGCCATTACAAG	GCCACTGAGAAGTGTCTAATTGA
Erg	CGTCCTCAGTTAGATCCTTACC	CTGTCTGACAGGAGTTCGAG
Myc	GCGACTCTGAAGAAGAGCAAGA	GACCTCTTGGCAGGGGTTTG
Pigp	ATGGTGGAAAATTCACCGTCG	ACGAAAGCCCACACAAGATAAA
Pold2	ACCTGTGGCAACTTACACCAA	GTGGCATAAATATGGGCGTACT
Tfrc	GTGAAACTGGCTGAAACGGAG	GGTCTGCCCAATATAAGCGAGA

Mathematical modelling of HSC and MPP behaviour

Generation of a model for HSC and MPP behaviour

We build on the model of Wang *et al.* where delay differential equations are used to model irradiation-induced differentiation of HSCs 0–24 h after gamma-irradiation (Wang et al. 2012). Since the time frame (our data spans 28 days) which we are interested in is different, we developed a simpler model where we focused on the effect of *Erg* in the proliferation or differentiation of HSCs and MPPs.

Experimental data points were averaged for each time point among all replicates. The data was then normalized against HSC $Erg^{fl/fl}$ time series, to account for the effect of pIpC injection. The normalized data is shown below.

Modelling 1: Experimental data normalized

Time (days)	*HSC $Erg^{fl/fl}$	*MPP $Erg^{fl/fl}$	*HSC $Erg^{\Delta/\Delta}$	*MPP $Erg^{\Delta/\Delta}$
0	0.036	0.127	0.046	0.150
14	0.036	0.117	0.001	0.023
28	0.036	0.135	0.008	0.025

*cell counts in million

The normalized experimental data for $Erg^{fl/fl}$ was used to fit the model of HSC/MPP behaviour. Optimal parameter values were found by minimising the squared error between the experimental data and the model. To improve the accuracy of the fitting for the $Erg^{\Delta/\Delta}$ data, an artificial time point was added at day 90, assuming that HSC and MPP levels have reached the levels seen at day zero

In order to model the behaviour of HSCs and MPPs we introduce the following parameters:

Modelling 2: Model parameters description

Parameter	**Description**	**Unit**
r_{wt}	growth rate of HSCs	day^{-1}
K	number of HSCs in normal BM	cells
d_{wt}	differentiation rate of HSCs	day^{-1}
a	acceleration of differentiation for $Erg^{\Delta/\Delta}$ HSCs	none
r_{mwt}	growth rate of MPPs	day^{-1}
M	number of MPPs in normal BM	cells
d_{mko}	differentiation/death rate of $Erg^{\Delta/\Delta}$ MPPs	day^{-1}
d_{mwt}	differentiation/death rate of $Erg^{fl/fl}$ MPPs	day^{-1}

Let us describe a population of HSCs which we call *H*. A first model assumes that these HSCs are produced by the BM in a certain fashion, *f*. The differential equation for the growth of a population of size *H(t)* as a function of continuous time *t* is expressed as follows:

$$\frac{d(H(t))}{dt} = f(H(t)) \qquad (1.1)$$

f is a function that relates the populations rate of change *dH(t)/dt*, to its size *H(t)* at the time *t*. After integration, we get *H*, the number of HSCs, as a function of time. The most straightforward model could be a model where HSCs are produced at a steady rate which is expressed as follows:

$$\frac{dH}{dt} = k \qquad (1.2)$$

This model depicts that the variation of H over time is equal to the constant k and could describe the amount of MPPs produced by the BM from HSCs. The expression of the solution is the following: $H(t)=H_0 + kt$, where H_0 is the initial amount of HSCs. Plotting the solution gives a straight line which intercepts the y-axis at H_0. It is assumed here that HSCs do not die. Therefore, in order to achieve a biologically relevant model, we introduce the term $-\delta_H H$ which describes the death of the HSCs, with δ_H being the death rate of HSCs *per day*. The model is consequently expressed as follows:

$$\frac{dH}{dt} = k - \delta_H H \tag{1.3}$$

The analytical solution $H(t)$ is expressed as follows:

$$H(t) = \frac{k}{\delta_H}(1 - e^{-\delta_H t}) - H_0 e^{-\delta_H t} \tag{1.4}$$

The negative term on the right represents the loss from the initial value. For the term to the left, when time goes to infinity, the exponential inside the parenthesis will approach zero, and the final value or *steady state* of the model will approach $\lim_{t\to\infty} H(t) = k/\delta_H$, which is a function of the growth rate k and the death rate δ_H. The expression for the death of the HSCs is called an exponential decay, and is a common way to describe many biological phenomena. When the population stabilizes, it is possible to assume that the variation is equal to zero which allows identification of steady states (equilibrium points) for the model. In the case of the model in equation 1.3, it is possible to find the steady state just by assuming that dH/dt is equal to zero which is solved with the following equation:

$$\frac{dH}{dt} = k - \delta_H H = 0 \tag{1.5}$$

This has the solution $\bar{H} = k/\delta_H$, where $\bar{H}$ denotes the equilibrium point at homeostasis for *Erg*$^{fl/fl}$ mice.

The models developed previously for growth can account simultaneously for both death and birth. Let μ and δ_H be the birth and death rates respectively, one can formulate a model of exponential growth

that combines both of these rates. Again, let H be the variable population, a reasonable model of exponential growth would be:

$$\frac{dH}{dt} = \mu H - \delta_H H = (\mu - \delta_H)H = rH \qquad (1.6)$$

where r equals $(\mu - \delta_H)$. If $r > 0$, the population will expand, and if $r < 0$, it will decrease and reach zero. In the case of $r > 0$, when t becomes very large, the population is infinitely large ($\lim_{t\to\infty} H(t) = \infty$). Biologically, there is no phenomenon that results in an infinitely large population, and there is often a limit above which a number of cells cannot be larger due to restrictions in space, nutrient availability *etc*. In order to reflect this behavior in mathematical terms, it is possible to rely on the classic model of logistic growth where the rate of reproduction is proportional to the existing population and the amount of available resources. This model is formalized by the following differential equation:

$$\frac{dH}{dt} = rH(1 - \frac{H}{K}) \qquad (1.7)$$

where the constant r defines the growth rate and K is called the carrying capacity. The left part of the expression of this model resembles the exponential growth model previously shown in equation 1.6 which forces it to behave similarly to an exponential model when H is small, resulting in an almost exponential expansion of the population. The expression on the right, $(1 - H/K)$, which we call A for simplicity, is a limiting term. If H becomes larger than K, A takes a negative value, and the growth becomes negative; in other words, the population decreases. Thus, the carrying capacity K is a limit above which the population cannot expand anymore. K is determined from HSC numbers in $Erg^{fl/fl}$ mice.

Let us assume a simple model of early hematopoiesis where HSCs (H) differentiate into MPPs *(M)*. The model assumes a constant production rate (σ) and an exponential decay for the HSCs $\delta_H H$. A proportion of the HSCs will differentiate into MPPs, at a rate $d_h H$. Similarly, we combine the loss of

MPPs with one term that combines both differentiation and apoptosis d_mM. This leads to the following model:

$$\begin{cases} \frac{dH}{dt} = \sigma - \delta_H H - d_h H \\ \frac{dM}{dt} = d_h H - d_m M \end{cases} \quad (1.8)$$

Since we know neither the production rate nor the death rate of HSCs but only the number of HSCs at equilibrium, we modify the HSC equation to use a logistic growth model. The number of HSCs in normal BM can be assumed to be between 35000 and 45000 HSCs (see experimental data in *Modelling 1*); for simplicity, we set K to 50000. The MPP population is expected to proliferate and hence we include a logistic growth term also for this population, with parameter r_m *and* K_m for growth and normal number of HSCs respectively. K_m is assumed to be 150000 (see experimental data in *Modelling 1*). The model is expressed as follows:

$$\begin{cases} \frac{dH}{dt} = rH(1-\frac{H}{K}) - d_h H \\ \frac{dM}{dt} = d_h H + r_m M(1-\frac{M}{K_m}) - d_m M \end{cases} \quad (1.9)$$

Modelling the behaviour of HSCs/MPPs following pIpC injections in $Erg^{fl/fl}$ *and* $Erg^{\Delta/\Delta}$ *BM*

Cre-mediated gene deletion in *in vivo* models may not target 100% of the cells. Therefore, there might be a small, albeit non-negligible amount of WT cells still present after pIpC injections. Whether these escaper $Erg^{fl/fl}$ cells have a competitive advantage or not depends on the relative replication, death or differentiation rates of those cells compared to their $Erg^{\Delta/\Delta}$ counterpart. The model in equation 1.9 is extended to take into account two kinds of HSCs and MPPs in pIpC treated $Erg^{fl/fl}$; *Mx1Cre* mice: those with $Erg^{fl/fl}$ and those with $Erg^{\Delta/\Delta}$ genotypes. They are referred to as H_wt, M_{wt}, H_{ko} and M_{ko}, respectively. To reflect that both kinds of HSCs and MPPs are competing for the same space in the BM niche, we modify the model slightly to take this limiting carrying capacity into account. To facilitate the

reading of the model, we also assume that the differentiation rates for $Erg^{\Delta/\Delta}$ HSCs is the same as those carrying $Erg^{fl/fl}$, but with a modifier parameter α. Experimental work showed (Figure 4) that proliferation and apoptosis were not affected by the loss of *Erg*. To reflect this, the growth rate in the logistic terms for HSCs and MPPs are set to be the same for $Erg^{\Delta/\Delta}$ and $Erg^{fl/fl}$ populations (r_{wt} for HSCs and r_{mwt} for MPPs). As $Erg^{\Delta/\Delta}$ and $Erg^{fl/fl}$ MPPs may differentiate at different rate, we introduce two separate terms for this process. This final model of early hematopoiesis with interaction and modified differentiation rates is expressed as follows (see also flowchart in Figure S4A):

$$\begin{cases} \frac{dH_{ko}}{dt} = r_{wt}H_{ko}(1-\frac{H_{ko}+H_{wt}}{K}) - \alpha d_{wt}H_{ko} \\ \frac{dM_{ko}}{dt} = d_{wt}H_{ko} + r_{mwt}M_{ko}(1-\frac{M_{ko}+M_{wt}}{M}) - d_{mko}M_{ko} \\ \frac{dH_{wt}}{dt} = r_{wt}H_{wt}(1-\frac{H_{ko}+H_{wt}}{K}) - d_{wt}H_{wt} \\ \frac{dM_{wt}}{dt} = d_{wt}H_{wt} + r_{mwt}M_{wt}(1-\frac{M_{ko}+M_{wt}}{M}) - d_{mwt}M_{wt} \end{cases} \quad (1.10)$$

Fitting model to data

$Erg^{fl/fl}$: The final model (eq. 1.10) is used to find initial parameter values for r_{wt}, d_{wt} and d_m. Initial conditions are taken from the $Erg^{fl/fl}$ data at day zero for $Erg^{fl/fl}$ HSCs and MPPs, while $Erg^{\Delta/\Delta}$ HSCs and MPPs are set to zero. *K* is set to 50000. The parameter values found were subsequently used to set starting guess-values for the fitting algorithm.

$Erg^{\Delta/\Delta}$: Let H_{tot} be the sum of H_{wt} and H_{ko}. Let M_{tot} be the sum of M_{wt} and M_{ko}. We fit the data from the $Erg^{\Delta/\Delta}$ experiments to H_{tot} and M_{tot} using the final model (eq. 1.10) (assuming 45000 $Erg^{\Delta/\Delta}$ and 5 $Erg^{fl/fl}$ HSCs; 157000 $Erg^{\Delta/\Delta}$ and zero $Erg^{fl/fl}$ MPPs to start with). First, the HSC data is used to compute the growth rate and differentiation rates for both populations of HSCs. To improve the fit, the MPP data is also fitted at the same time; the differentiation rate for $Erg^{fl/fl}$ HSCs, proliferation rate and differentiation/death rates of MPPs are computed. The resulting parameter values can be found in

Modelling 3 (see also Figure S4B for a plot of the model). We find that $Erg^{\Delta/\Delta}$ HSCs differentiate 24 times faster than their $Erg^{fl/fl}$ counterparts. By increasing the value of α, the model still reach acceptable fit errors, but the HSC populations decrease very fast which is not what is observed in experiments. Experiments show that after three weeks, the HSCs that are found in pIpC treated *Erg* $^{fl/fl}$;*Mx1Cre* mice are mainly $Erg^{\Delta/\Delta}$ (Figure S4C). Therefore, we conducted experiments where the value of α was increased to find the time point where $Erg^{\Delta/\Delta}$ HSCs would be predicted to disappear. We report the HSC counts at day 20 and observe that with an α–value of around 25, the $Erg^{\Delta/\Delta}$ HSCs are gone 30 days after pIpC injections, and that there is a majority of $Erg^{fl/fl}$ HSCs. Values of α below that level prevent the model from fitting properly while values above predict an $Erg^{\Delta/\Delta}$ HSC population crash, which do not reflect the experiments conducted. We therefore conclude that there is a small window of optimal α values that allow the model to fit the data while being in agreement with the experimental data (Figure S4C, D). The model also shows that the proliferation of HSCs is not constant over time consistent with the known effects of injecting pIpC (Figure S4E). To test the robustness of the model when varying the relative proliferation rates of $Erg^{fl/fl}$ and $Erg^{\Delta/\Delta}$ HSCs, the proliferation rate of $Erg^{\Delta/\Delta}$ HSCs is artificially increased two-fold and the model is then fitted to the data again, using varying values of α. With an optimal value of 32.5 the model can again be fitted, but the $Erg^{\Delta/\Delta}$ HSCs disappear completely after 20 days which is not in agreement with experimental data. Halving the proliferation rate of $Erg^{\Delta/\Delta}$ HSCs and fitting the model again results in a value for α = 17.0 which we deem an acceptable fit-error. This shows that even if the proliferation rate of $Erg^{\Delta/\Delta}$ HSCs may be increased or decreased compared to $Erg^{fl/fl}$ HSCs, the only way for the model to fit the data is by increasing the differentiation rate of $Erg^{\Delta/\Delta}$ HSCs (Figure S4F). Overall, our model implicates that $Erg^{\Delta/\Delta}$ HSCs differentiate to MPPs faster than their $Erg^{fl/fl}$ counterparts.

Modelling 3: Parameter values for best fitted model

Parameter	**Value for best model fit**
r_{wt}	0.3202
K	50000
d_{wt}	0.025
α	24.3162
r_{mwt}	0.43
M	150000
d_{mko}	0.4709
d_{mwt}	0.005

Supplemental References

Breems DA, Blokland EA, Neben S, Ploemacher RE. 1994. Frequency analysis of human primitive haematopoietic stem cell subsets using a cobblestone area forming cell assay. *Leukemia* **8**: 1095-1104.

de Boer J, Williams A, Skavdis G, Harker N, Coles M, Tolaini M, Norton T, Williams K, Roderick K, Potocnik AJ et al. 2003. Transgenic mice with hematopoietic and lymphoid specific expression of Cre. *European journal of immunology* **33**: 314-325.

Dymecki SM. 1996. Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 6191-6196.

Kuhn R, Schwenk F, Aguet M, Rajewsky K. 1995. Inducible gene targeting in mice. *Science* **269**: 1427-1429.

Okabe M, Ikawa M, Kominami K, Nakanishi T, Nishimune Y. 1997. 'Green mice' as a source of ubiquitous green cells. *FEBS letters* **407**: 313-319.

Ploemacher RE, van der Sluijs JP, Voerman JS, Brons NH. 1989. An in vitro limiting-dilution assay of long-term repopulating hematopoietic stem cells in the mouse. *Blood* **74**: 2755-2763.

Schwenk F, Baron U, Rajewsky K. 1995. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic acids research* **23**: 5080-5081.

Ventura A, Kirsch DG, McLaughlin ME, Tuveson DA, Grimm J, Lintault L, Newman J, Reczek EE, Weissleder R, Jacks T. 2007. Restoration of p53 function leads to tumour regression in vivo. *Nature* **445**: 661-665.

Wang J, Sun Q, Morita Y, Jiang H, Gross A, Lechel A, Hildner K, Guachalla LM, Gompf A, Hartmann D et al. 2012. A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage. *Cell* **148**: 1001-1014.

Weischenfeldt J, Damgaard I, Bryder D, Theilgaard-Monch K, Thoren LA, Nielsen FC, Jacobsen SE, Nerlov C, Porse BT. 2008. NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. *Genes Dev* **22**: 1381-1396.

Zhang P, Li MZ, Elledge SJ. 2002. Towards genetic genome projects: genomic library screening and gene-targeting vector construction in a single step. *Nature genetics* **30**: 31-39.